Peer-review report of

Franco, V.R., Barros, G.W.G., & Jiménez, M. (2024). A generalized approach for Bayesian Gaussian graphical models. *advances.in/psychology, 2*, e533499. https://doi.org/10.56296/aip00022

Round 1

Dear authors, it is my pleasure to share the two detailed reviews we received about your paper on the “Generalized Approach for Bayesian Gaussian Graphical Models”.

The two reviews critically assessed your manuscript introducing a novel approach to Bayesian Gaussian Graphical Models (BGGMs) using a transformation of Cholesky decomposition matrix values. This method aims to enhance the flexibility and applicability of BGGMs by allowing for customizable priors and the estimation of regularized partial correlations across variable types. The manuscript also presents the gbggm R package, designed to implement this novel approach.

Reviewer 1 offers a detailed critique, acknowledging the potential value of the paper but ***highlighting significant issues that detract from its readability and scholarly contribution***.

Key criticisms include:

- Lack of clarity and detail in presenting the proposed approach, making it difficult for others to adapt.

- Structural and grammatical deficiencies in the manuscript, necessitating a substantial rewrite.

- A toy example that is unconvincing and inadequate for demonstrating the method's applicability, particularly for psychological research.

- The need for a more thorough literature review, especially on Bayesian shrinkage estimators and prior work on Cholesky decomposition.

- Concerns over the paper's novelty and its comparison with existing methods, including the use of Wishart priors and the handling of ordinal data.

- Suggestions for improving the clarity of methodological explanations and the discussion of potential applications and limitations.

Reviewer 2 raises concerns about the originality and specificity of the contribution, questioning whether the use of Cholesky decomposition for Bayesian estimation of covariance matrices represents a significant advancement. The review points out:

- Ambiguities in the paper's claims and a lack of clear differentiation from existing methodologies.

- The absence of evaluation using widely accepted estimation methods like Hamiltonian MCMC or variational inference.

- Requests for more explicit explanations and justifications of the methodological choices, including the handling of high-dimensional parameter spaces and the rationale behind specific Bayesian priors.

- The need for a clearer demonstration of the method's advantages, particularly in handling binary data and specifying latent variable models.

- Suggestions for incorporating modern Bayesian estimation techniques and providing more detailed simulation data to validate the proposed approach.

Both reviews express interest in the proposed method but emphasize the need for major revisions to clarify the contributions, improve the manuscript's structure and exposition, and more effectively demonstrate the method's utility and novelty.

I must admit that the reviewers are quite critical of your work. One reviewer expresses significant concerns about the originality and clarity of the paper's contribution, specifically questioning the novelty of using Cholesky decomposition for estimating covariance/correlation matrices, a technique already well-established in the field. The reviewer criticizes the omission of main Bayesian estimation methods, which undermines the strength of the provided toy example. The critique highlights a lack of necessary detail throughout the paper, suggesting that it falls short of the standards expected for submission. While the reviewer sees potential for improvement if the authors could more clearly define their contribution, there is skepticism regarding the paper's ability to present a truly novel approach.

Despite the critical feedback, I am willing to offer you an opportunity to revise your manuscript. If you can thoroughly address the concerns raised by the reviewers and effectively persuade both them and me with your revisions, there remains a possibility for your paper to be accepted. Please be aware that this will be a substantial undertaking, akin to a Herculean task. However, if you are prepared to dedicate the required effort to resolve all identified issues, I am open to reconsidering your submission for publication.

Reviewer 1:

The current paper proposes a generalized approach to estimating Gaussian graphical models in a Bayesian framework, using the Cholesky decomposition. I have a number of concerns with this paper that I detail in the comments below. My main concern is that it is unclear what the authors are claiming to have developed? Are they claiming to have developed the Cholesky decomposition of covariance matrices for ease of estimation in a Bayesian framework? I hope not, as this would ignore the fact that these sorts of decompositions have been used in Bayesian analysis to impose priors on covariance/correlation matrices for years. For example, Stan, a major Bayesian estimation software (well, a probabilistic programming language), has detailed the Cholesky transformation in their documentation since

I've been working with Stan (8+ years). Furthermore, the authors do not evaluate estimating these models with the most widely used estimation methods, ala, Hamiltonian MCMC or variational inference. So, the paper would benefit greatly from clarifying what the specific contribution is.

1. On page 6, the approach of Williams 2020 is said to be "equivalent to regularized Bayesian GGMS." I don't quite see the equivalence in the description of the approach. The use of the diagonal hyperparameter matrix seems to be more equivalent to a ridge type regularization, while the use of post-hoc tests for removing GGMs is entirely outside of what I'd consider regularization. More clarification on this statement would be helpful to place the previous work in the rest of the literature.

2. Page 7: Why would the MCMC sampler have an issue with sampling the entire matrix at once? Is this related to the curse of dimensionality? And how would modern MCMC sampler such as NUTS perform, as they are built to handle the sampling of high dimensional parameter spaces?

3. Page 7-8: What is MAP estimation? Unless I am mistaken, MAPs are calculated from the posterior samples, and act as a point estimate, similar to EAPs. Additionally while I agree that the MAP will likely have issues with producing a positive semidefinite matrix (as, what is the mode of a sample of matrices?), the EAP will produce a positive definite matrix as the sum of positive definite matrices is indeed positive definite.

4. Page 9: The definitions of ridge and LASSO using the Bayesian priors should be directly cited (I don't know if they come from the Erp, 2020 reference). Or, if this is original work, the authors should put the proofs in an appendix.

5. Page 10: How does the authors' method allow for simultaneous estimation of the thresholds?

6. Page 16: What is being shown in the table? Bayesian samplers return a posterior distribution of the parameter values, and point estimates need to be derived. So, are these values the EAP? The MAP? The last value pulled by the sampler? Additionally, the benefit of a Bayesian approach is in the quantification of uncertainty, so where are the credible intervals?

7. The plots would benefit from a vertical line picking out the true value.

8. General: The extensions proposed are quite interesting, but the paper would benefit if the authors described how, and then showed a toy simulation demonstrating that the method works. As it stands right now, this section should be in the discussion under Future Directions.

9. General: As I mentioned in the opening paragraph, I am concerned that something like Stan, with its implementation of Hamiltonian MCMC and also variational inference, was not considered here. I'd strongly suggest the authors evaluate their approach using these modern methods.

10. Page 13: I am quite confused about the data simulation procedure: It seems as though the data is simulated as binary, but by dichotomizing a sampled normal distribution. How is that dichotomization occurring? What is the threshold. More concerningly, how is the proposed method handling the binary nature of the data? It seems as though the data is being treated as continuous in the estimation (ala, just fed into the sampler), when the nice thing about a Bayesian framework is that one can easily specify the latent variable and measurement model for the binary data. Indeed, even the dichotomization thresholds can be estimated.

In closing, I found this paper to be interesting, but in need of more clarification and details throughout. A generalized method for estimating BGGMS would be quite useful, and I hope that more detail would make the contribution of this paper clear.

Reviewer 2:

This paper introduces a novel approach to Bayesian Gaussian Graphical Models (BGGMs) to address limitations in existing models, particularly their inflexibility in accommodating various priors and likelihoods. The proposed method utilizes a transformation of Cholesky decomposition matrix values as parameters, allowing for the estimation of regularized partial correlations between diverse variable types with customizable priors. The gbggm R package has been developed to implement models based on this approach, showcasing its potential through a toy simulation, thereby broadening the applicability and flexibility of BGGMs in different contexts.

We thank the authors for their interesting manuscript and the opportunity to read and review it. Bayesian graphical modeling is a growing field in psychology and we would consider a general approach to the adaptation of models to be very fruitful. However, there are several critical issues with the manuscript that significantly impede its readability and scholarly value. First, we fear that the introduction of the proposed approach lacks sufficient clarity and detail to allow other researchers to adapt it.

Second, the toy example could be more convincing. As of now, it is presented on a network that is smaller than most psychological applications and on a method (ordinal data) where the proposed approach is not needed. Third, the writing is unstructured, verbose, and full of grammatical, spelling, and punctuation errors. The reading flow of the paper could be improved.

In conclusion, despite the very interesting idea at the core of the paper, the paper would need a major revision to make its benefit to the field pronounced. We elaborate our concerns in detail below and hope they will be helpful for the authors.

Major comments:

The writing suffers from structural deficiencies, poor organization, and frequent grammatical errors, spelling mistakes, and missing punctuation. The authors should undertake a substantial rewrite to enhance the manuscript's clarity and cohesiveness.

We suggest to reconsider the notation for the covariance matrix (currently denoted with R). R typically denotes the correlation matrix, and as such the notation might lead to confusion. We suggest to use the common notation for the covariance matrix \Sigma.

The paper highlights the inability of BGGM to test for the presence or absence of edges. However, Williams et al. use the Savage-Dickey ratio to obtain an inclusion Bayes factor. Furthermore, the authors state that Williams's approach is tailored for continuous data, but the approach also allows for ordinal, mixed, and binary data through a multivariate probit or copula transformation. Therefore, it is not clear whether and to what extent to discussed downsides still hold.

We suggest to include further (and in our opinion crucial) literature on Bayesian shrinkage estimators for GGMs, such as Li et al. (2019), Wang (2012), and Sagar (2022), which would significantly enrich the paper's theoretical framework. (Li et al., (2019): The Graphical Horseshoe Estimator for Inverse Covariance Matrices; Wang (2012): "Bayesian Graphical Lasso Models and Efficient Posterior Computation"; Sagar (2022): "Precision Matrix Estimation under the Horseshoe-like Prior–Penalty Dual".)

Sampling an entire matrix is extremely desirable for the quality of the MCMC samples, if it doesn't kill the acceptance probability (e.g., the part about Wishart

distributions is questionable at best). Could authors elaborate why they suggest that this is a concern?

Using the cholesky decomposition to model the covariance matrix is not new. This is most certainly not a problem as authors introduce it as a general modeling approach. Nonetheless, it would be necessary to discuss previous adaptations such as (Roverato (2000): "Cholesky decomposition of a hyper inverse Wishart matrix").

Introducing the Cholesky decomposition is the pivotal aspect of the paper. We suggest to increase its clarity and readers comprehension with detail and with structure.

It would be useful to compare this prior setup with the commonly used Wishart priors. A major difference between the Wishart prior and the prior proposed by the authors is that they specify independent prior distributions for each element of the Cholesky (precision) matrix. By setting these priors independently, it is not possible to account for scaling effects that may occur when considering networks of different sizes.

The section on possible adaptations could be made more clear. In the current version, it is not clear to us (and we fear other readers might have a similar concern), how the direct adaptation of the approach to the examples should be conducted. Focusing on a specific adaptation and discussing it in detail might be more fruitful to readers.

The last paragraph of the possible adaptations highlights a drawback of the approach. We would suggest taking it out from this section and rather adding to a limitations section in the discussion.

We were happy to see an example in the paper. Unfortunately, we do not find the it convincing; it is inadequate and possesses too few variables ($p = 3$). Li et al. showcased a simulation of $p = 100$ and $p = 200$. While that might be too extensive, we suggest that the authors at least consider an example with $p = 7$. Furthermore in the toy example, we suggest that authors explicitly specify both the prior and likelihood.

The toy example is not necessarily as convincing as one could imagine with the proposed method. The ordinal model does not mandate positive definiteness and thus the Cholesky decomposition need not exist. We fear that this makes the

comparison of the proposed work to some extent meaningless. Do authors propose to treat ordinal data as continuous? Furthermore, all methods introduced in the paper are available to model binary data (mostly through directly modeling it as an Ising model or even through a copula transformation). We would suggest the authors find a more convincing example.

In the discussion section, several huge drawbacks are discussed. We would suggest making the conclusion paragraph stronger to highlight the benefits of the current approach, despite the drawbacks.

Minor issues:

Line 45-46: The line “as the variables are conditionally independent on each other” does not make much sense. You could say that when a partial correlation between two variables is zero, these variables are conditionally independent given the remaining variables in the network. The reference to Epskamp et al. does not fit this context as it discusses the relation between binary network and latent variable models. You could use:

https://www.tandfonline.com/doi/full/10.1080/00273171.2018.1454823

Line 60 - 65: It is strange to discuss social network models as the origin of graphical models in psychology. There is, of course, a long tradition of graphical modeling, but that is entirely unrelated to the literature on social network models.

Line 95–97: There are hundreds of papers on Bayesian graphical modeling, I suspect, but not so many (yet) in psychology. There are some missing here though, especially to the work of Donald Williams and Maarten Marsman, but see also work such as: https://arxiv.org/abs/1911.07142.

missing references (line 171) or lack of citations for statements (line 59 +60,) and wrong citation style (line 111)

text overlap in figures

Spelling errors (e.g., “intuitive” in the abstract)

Style: e.g., line 55 “In psychological research, research into”

Line 123: It is unclear what is meant here with “a density of 1.0”. Do the authors mean a probability mass of one like in a discrete spike and slab prior? Where is this mass placed? Perhaps the authors could explain this a bit more.

Line 343, caption Table 1: Halsberg -> Haslbeck.

Round 2:

Dear Authors,

Thank you for submitting your revised manuscript "A Generalized Approach for Bayesian Gaussian Graphical Models" for consideration in the Special Issue on Network Methods in Advances in Psychology.

We are pleased to inform you that your paper has been accepted for publication. The reviewers and editors found that you have addressed the previous comments and concerns thoroughly, resulting in a significantly improved manuscript. In particular, we appreciate the following changes:

1. The substantial rewrite to enhance clarity and cohesiveness throughout the manuscript.
2. The inclusion of additional relevant literature, especially on Bayesian shrinkage estimators for Gaussian Graphical Models.
3. The improved explanation of the Cholesky decomposition and its role in your proposed method.
4. The updated toy simulation with 7 variables, providing a more realistic demonstration of your approach.
5. The addition of an empirical example comparing your method to existing approaches and implementing it using various estimation methods, including Stan.
6. The clarification of your method's novelty and its positioning within the existing literature on Bayesian Gaussian Graphical Models.
7. The revised discussion of possible adaptations and limitations of your approach.
8. The overall improvements in notation, references, and technical explanations throughout the paper.

We believe these changes have significantly strengthened your manuscript and made its contribution to the field clearer and more compelling. Your generalized

approach for Bayesian Gaussian Graphical Models will be a valuable addition to the literature on network methods in psychology.

Our editorial office will be in touch if any additional materials required for publication.

Congratulations on your excellent work, and thank you for your contribution to Advances in Psychology.

Sincerely,

Hudson Golino, Editor, Advances in Psychology

Reviewer 1:

The authors have addressed my comments, and improved the manuscript significantly!